# 情境特征及其在情感分类模型中的应用 *

刘　栋[a]，张彩环[b]

(洛阳师范学院 a. 信息技术学院; b. 数学科学学院, 河南 洛阳 471022)

**摘　要：**为研究情境特征在文本分类中的作用，提出了一种层级双向 LSTM 模型用于情感分类问题。该模型首先将句子分词，把词向量作为第一层双向 LSTM 模型的输入；其次从文档中提取出稠密、连续的向量作为情境特征；然后将第一层模型的输出向量和情境向量共同输入第二层双向 LSTM；最后将这种层级双向的 LSTM 模型的输出向量通过 sigmoid 函数进行分类。情境向量作用于每个句子，一致的情感得到增强，不一致的情感被弱化，从而提高了分类的精度。在两个公开数据集上的实验表明，整合了情境特征的层级双向 LSTM 取得较优的精度。除此之外，通过在一个包含 2 万余条中文评论的公开数据集上对模型进行测试，表明该模型测试正确率相比于普通的 LSTM 和双向 LSTM 都有提升，说明情境特征对于提升情感分类的作用比较显著。

**关键词：**情境特征；情感分类；层级双向 LSTM

**中图分类号：**TP391　　doi: 10.19734/j.issn.1001-3695.2018.06.0459

# Context feature and its application on sentiment classification model

Liu Dong[a], Zhang Caihuan[b]

(*a. School of Information Technology*, *b. School of Mathematics*, *Luoyang Normal University*, *Luoyang Henan* 471022, *China*)

**Abstract:** Focusing on the effect of context feature on sentiment classification problem, this paper presented a hierarchical Bi-LSTM sentiment classification model. The model includes two level Bi-LSTMs. The model accepted a series of word embedding as input in sentence-level. The context feature of whole document was transformed into a dense and continuous vector, concatenated with output vectors of the sentence-level Bi-LSTMs and fed into sigmoid layer to predict sentiment polarity of the document. Intuitively, the model computes a context vector mask for every sentence in document. The consistence polarity was enhanced, and the inconsistent was weakened. This resulted in better accuracy. The experiment on IMDB and Yelp 2013 dataset shows that this hierarchical model with context feature outperforms the common LSTM and Bi-LSTM models significantly, and is competitive with the state-of-the-art results. Experiment on another dataset including 20 thousands Chinese reviews was used to analyze the causes of this competitive advantage.

**Key words:** context feature; sentiment classification; hierarchical Bi-LSTM



## 0　引言

文本分析的一个重要研究方向是对文本所表达的情感进行分类。当前，随着互联网以及移动端应用的发展和普及，网络上形成了海量的文本，特别是评论，表达了用户对商品、服务、电影及时政等的观点和看法。对这些文本进行深入的分析和挖掘有着重要的应用价值，并产生巨大的经济效益[1]，比如通过分析商品评论能够为商家提供产品反馈，进一步改进服务质量，优化营销策略；对网络评论进行情感分析，及时了解社情民意等。因此对文本的情感分类进行研究是十分有意义的。

文本情感分类从类别上来分，大多数研究都分成正向和负向两种极性；也有涉及到三分类（正向、负向、中立）[3]和多分类的（将用户的情感按强度量化为多个等级）[4]。从所处理的文本粒度来看，可以分为文档级、句子级和词语级[5]。从所处理的语言来看，可以分为单语言和跨语言情感分类。从涉及的领域来看，又可以分为单领域及跨领域。

从文本情感分类技术上来看，早期的研究主要依赖于情感词汇，通过从文本中挑选出情感词作为特征，再利用这些特征进行情感分类，这类方法的准确率并不理想[2]。之后机器学习的方法包括有监督、无监督、半监督学习的方法被应用于文本情感分类问题使得分类结果得以提升[6~17]。随着深度学习技术在图像和语音方面所取得的巨大成功，人们将此类技术运用于

**收稿日期：**2018-06-19；**修回日期：**2018-08-11　**基金项目：**政府间国际科技创新合作重点专项资助项目（2016YFE0104600）；河南省科技攻关资助项目（182102210434）；河南省高等学校重点科研项目（15A520024）

**作者简介：**刘栋（1971-），男，河南博爱人，讲师，博士，主要研究方向为信息检索、云计算、自然语言处理、数据挖掘等（212616@qq.com）；张彩环（1969-），女，河南新郑人，副教授，博士，主要研究方向为计算数学，组合数学等.

情感分类问题，也取得了一定的进展，Severyn 等人[18]使用无监督的语言模型来训练词向量，这些词向量作为一个深度 CNN（convolutional neural network）的初始化参数来对 Twitter 进行情感分类，在 Semeval-2015 的两个子任务中取得很好的结果；Tang 等人[19]将目标信息整合到 LSTM 中，建模了目标与它上下文之间的相关性，提升了情感分类的精度；Nguyen 等人[20]提出一种基于 RNN（recusive neural network）情感分类方法，该方法考虑了依赖和句子的成分组成树。这类方法取得的结果说明深度学习在自然语言理解方面大有可为。然而，这些方法只考虑了短语级，或者句子级的关系，不能建模句子与句子间的情感延续、加强或转折等关系。深度神经网络能够获得不同抽象级别的特征，因此增加神经网络层数无疑是一种重要的方法，但是对于自然语言理解来说，增加神经网络的层数对计算资源和数据集的大小都产生了新的要求，相比来说，使用重复堆叠的层级模型更加合适，比如 Denil 等人[21]就使用了一个层级的卷积神经网络来抽取评论中的重要句子；Li 等人[22]将层级 LSTM（long short term memory）自动编码器应用于段落生成和重构；Ji 等人[23]基于 RNN（recurrent neural networks）和 CNN 进行短文本的分类，但该方法是句子级的，并且只考虑了目标前后有限个的句子；与本文采用方法最类似的是 Ruder 等人[24]的层级模型，他们同样采用了层级的、双向的 LSTM 模型进行情感分类，但不同的是，他们在句子级加入了产品方面的信息，而本文则是在每个句子中加入了情境信息。

在讨论文档所表示的情感时离不开它的上下文，也就是句子所处在情境，因此在情感分类的模型中加入文档的情境无疑是有益的。本文讨论了文档情境的分布式表示，这种表示方法曾经被Choi 等人[27]用在神经机器翻译模型中，取得了不错的效果。Ghosh 等人[31]在模型中整合入了情境特征（主题），在三个自然语言处理（NLP）任务上评估 CLSTM 模型：接续词语预测、接续语句预测、语句话题预测，能做到 20% 的提升。

本文的贡献和创新之处体现在如下几个方面：a)本文将情境信息表示为一个连续的稠密向量应用于句子的情感表达，与全局情境一致的情感得到了加强，不一致的情感被弱化，因此得到了更好的分类效果；b)通过实验说明了情感分类模型在加入了情境信息后，分类精度不仅优于普通的双向 LSTM 模型，而且与一些采用注意力机制的模型相比，在经过适当设置超参数后测试精度也有相当大的提高。

## 1　情境的表示及情感分类模型

### 1.1　情境的分布式表示

将每条评论看成是一个文档（document），每个文档截断为包含 ℓ 个句子（根据语料选择 10），如果不足 ℓ 个句子的，添加只包含空字符串的句子补全；每个句子通过分词，截断成包含$m$个词（5），不足$m$个的添加空字符串补全。这样第 $i$ 句子可以由词向量连接起来，表示为

$$s_i = (x_1^i, x_2^i, \cdots, x_m^i)$$

其中：$x_i \in \mathbb{R}^k$ 是第 i 个词的 k 维词向量。

同理，文档第 $i$ 个文档可以由句子连接起来，表示为

$$w_i = (s_1^i, s_2^i, \cdots, s_l^i)$$

在分析一条评论的情感时，可以把评论认为是一篇文档，整个文档构建了一个情境，因此情境可以表示为文档中所有的词组成的无序集合，同 Choi 等人[27]采用的方法一样，本文把情境$c^x$定义如下：

$$c^x = \frac{1}{T}\sum_{t=1}^{T} x_t$$

其中：$T$ 为文档中词的个数。

### 1.2　层级双向 LSTM 模型

LSTM 是一种特殊类型的 RNN，可以学习长期依赖信息[25]。LSTM 通过在 RNN 单元增加输入、输出和忘记门来建模长期依赖，在很多问题，LSTM 都取得了巨大的成功，并得到了广泛的使用。在对评论的情感分析中，人们认识到长期依赖问题同样有着重要的影响，一条评论不仅前面的句子可以反映出客户所要表达的感情，后面的句子很多时候或者是对情感的加强，或者是情感的反转，因此在模型中使用了层级双向的 LSTM，分为句子级和文档级。

a)在句子级设置双向 LSTM，对于句子中的词$w_t$，将$w_t$对应的词向量$x_t$、前一个单元的输出$h_{t-1}$ 和单元状态 $C_{t-1}$ 输入前向 LSTM 来计算下一个输出$h_t$和单元状态 $C_t$。先计算输入门 $i_t$ 和单元状态候选值 $C_t^{'}$ 以及忘记门 $f_t$：

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$
$$C_t^{'} = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$
$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

由此计算新单元状态值：

$$C_t = i_t * C_t^{'} + f_t * C_{t-1}$$

最后计算出输入门的值和本 LSTM 单元的输出 $h_t$：

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$
$$h_t = o_t * \tanh(C_t)$$

其中：$W_i, W_c, W_f, W_o, U_i, U_c, U_f, U_o$ 为权重矩阵，$b_i, b_c, b_f, b_o$ 为偏置向量。

这个前向 LSTM 是按照词出现的先后次序进行处理的，因此可以充分利用先前的信息；同理，后向 LSTM 以相反的次序处理，可以充分利用后面的信息。某时刻的输出由两个 LSTM 对应时刻的相应状态连接而成。

b)在句子级上面堆叠一个双向 LSTM。不同的是，这个双向 LSTM 是把句子级 LSTM 的输出与情境向量 $E$ 连接起来作为输入的，构成了第二层的双向 LSTM，增加了情境向量后的输入门 $i_t$ 和单元状态候选值 $C_t^{'}$ 以及忘记门 $f_t$ 修改为如下公式：

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + W_e E + b_i)$$
$$C_t^{'} = \tanh(W_c x_t + U_c h_{t-1} + W_e E + b_c)$$
$$f_t = \sigma(W_f x_t + U_f h_{t-1} + W_e E + b_f)$$
$$C_t = i_t C_t^{'} + f_t C_{t-1}$$
$$O_t = \sigma(W_o x_t + U_o h_{t-1} + W_e E + b_o)$$
$$h_t = O_t \tanh(C_t)$$

c)第二层 LSTM 的输出连接起来形成的向量被传递给一个 sigmoid 函数，产生该文档在情感分类上的概率分布。模型的结构如图 1 所示。

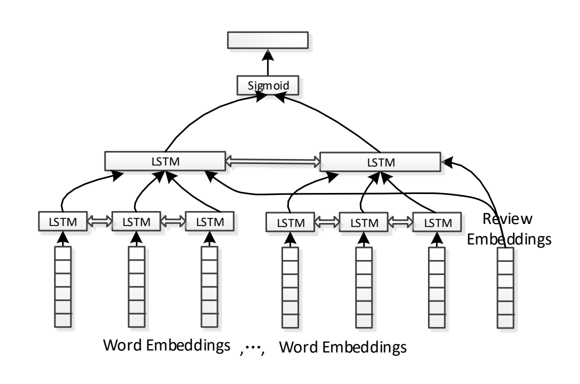



图 1　HB-LSTM 模型结构图

Fig.1　Structure of HB-LSTM

## 2　实验及结果分析

### 2.1　数据集与训练

本文在两个数据集上进行了评估：IMDB 和 Yelp 2013。数据集由 Tang 等人[30]建立并进行了相应的处理，训练集、开发集、测试集的比例为：8:1:1，与 Chen 等人[29]使用了相同的数据集，不同的是，在训练中 Chen 等使用了 200 维的预训练词向量，而本文是在模型中增加了嵌入层，生成了 256 维的词向量。基本的统计信息如下表所示：

表 1　数据集统计

Table 1　Statistics of datasets

| 数据集 | 类别数 | 文档数 | 句子/文档 | 词/句子 |
|---|---|---|---|---|
| IMDB | 10 | 84919 | 16.08 | 24.54 |
| Yelp2013 | 5 | 78966 | 10.89 | 17.38 |

实验中采用的基线模型是 Bi-LSTM，评价指标为测试精度（accuracy），模型的训练目标是最小化交叉熵函数（binary_crossentropy）,使用 Adam 优化器，mini-batches 设为 128。

为了便于对实验结果进行分析，将模型应用于另一个中文数据集[26]，数据集涉及 6 个领域，包含了 2 万多条评论。主要是商品及酒店等服务评论，最短的评论只有一个字，如“好”，“赞”等，最长的评论 2968 个字，是关于旅游和酒店服务投诉的。所有的评论都被标注了正面或者负面情感。经过打乱次序后取前 15000 个作为训练集，其余的作为测试集。

在模型的输入层，使用了 256 维的词向量，情景向量 256 维；LSTM 隐藏节点 128 个，Dropout 设为 0.5。在训练前所有文本进行了分词，使用的分词软件是 jieba 分词。

### 2.2　结果与比较分析

之前在数据集 IMDB 和 Yelp 2013 上进行了测试，并且与单层的 LSTM、Bi-LSTM 进行了比较，通过情境向量的作用，分类正确率有了一定的提升。本次实验主要是以 Bi-LSTM 作为基线，验证情境向量的作用，因此将之与带情境特征的 Bi-LSTM 和带情境特征的 HB-LSTM 两个模型进行比较，结果如表 2 所示，与之前一样采用了提前停机（early_stopping）的策略，容忍度(patience)为 2 个 epoch。

表 2　在 IMDB 和 Yelp2013 上的结果

Table 2　Results on IMDB and Yelp2013

| 模型 | IMDB | | Yelp 2013 | |
|---|---|---|---|---|
| | 训练精度 | 测试精度 | 训练精度 | 测试精度 |
| Bi-LSTM | 91.67 | 90.04 | 82.44 | 80.13 |
| Bi-LSTM with Context | 91.68 | **90.21** | **82.40** | **80.49** |
| HB-LSTM with Contex | 91.63 | **90.32** | **82.53** | **80.94** |

从表 2 可以看出，和 Bi-LSTM 相比，HB-LSTM 在两个数据集上的测试精度均有不同程度的提升。

由于使用了与 Chen 等人[29]使用的相同数据集，并且同样是情感分类，因此可以与该结果进行比较分析。Chen 等人实验中选择了 RMSE 作为损失函数，而本文采用的是交叉熵函数。表 3 中显示在这两种情况 HB-LSTM 与采用了注意力机制的 NSC 模型（NSC+UPA 表示 neural sentiment classification with user and products information）的测试精度比较。

表 3 两种损失函数情况下 HB-LSTM 与 NSC+UPA 的比较

Table 3　Effect of HB-LSTM and NSC+UPA with RMSE or binary_crossentropy as loss function.

| 模型 | IMDB | | Yelp 2013 | |
|---|---|---|---|---|
| | RMSE | Binary Crossentropy | RMSE | Binary Crossentropy |
| NSC+UPA | **53.3** | N/A | **65.0** | N/A |
| HB-LSTM | 50.03 | **90.32** | 48.78 | **80.94** |

从表 3 中可以看到结果的差异非常大，同样采用 RMSE 作为损失函数，HB-LSTM 的精度要比较 NSC+UPA 差 3~7 个百分点。但如果采用 Binary_Crossentropy 作为损失函数，精度几乎提升一倍，其中的原因有可能是一些超参数的设置不同，这还有待于进一步研究。

中文数据集是有关商品及服务评论的情感数据集，属于二分类问题。LSTM 模型在这样的二分类问题上已经取得了不错的效果。为了验证情境向量能否进一步改进模型的分类效果，实验选取单层 LSTM 作为基线模型，分别与双向 LSTM（Bi-LSTM）、层级双向 LSTM（HB-LSTM）进行比较。实验结果如表 4 所示。

表 4　中文评论数据集上的结果比较

Table 4　Experimental result and comparision

| 模型 | 训练精度 | 测试精度 |
|---|---|---|
| LSTM | 96.60 | 89.31 |
| Bi-LSTM | 96.62 | 89.45 |
| HBLSTM | **98.87** | **90.75** |

如表 4 所示，采用 LSTM 了的模型总体表现良好，LSTM 对邻近的信息进行了整合，较好地考虑了每个句子中前后一定范围内的依赖关系，双向 LSTM 的改善空间不大。但是增加了

情境特征后的层级双向 LSTM（HB-LSTM）正确率则有着明显的提升，实验结果说明，增加了情境向量后，全局的情境信息在特征中产生了作用，取得了优于基线模型的改善。表 5 中列出的几个例子都是 LSTM 错分的，用于对比 LSTM 和 HB-LSTM。

表 5　LSTM 错分的例子

Table 5　Some wrong classification example by LSTM

| 评论 | LSTM | HBLSTM |
|---|---|---|
| 很不错，到货很快，当天给客服打电话，下午安装人员就上门了，速度啊!但但当时没有水，没有试试，应该没事的，万一有问题还可以找厂家维护人员 | negative | positive |
| 原装用的气的热水器坏了,，换上了,，好用多，不会再出现要等 40 秒以上才会有热水，且还不稳定的情况了,，现在用上美的电热水器 5 秒后就出热水，且稳定的恒温，好用,，建议千万不要去买烧气的,!! 就美的电热水器，经济且很实用! | negative | positive |
| 网速快，且方便 VPN 连到公司内部网，IT 负责人把网络设置得不错，表扬一下。但前台服务员态度达不到满意的标准。酒店没有熨斗，因要见客户，只得拜托熨一件衬衫，说了 5 块，结帐时却要价 8 块，说是服务费和加急费要 3 块。打开玻璃淋浴间的门，才发觉右手至少要有 2 米长才可够得着浴巾，只好滴着水出来拿。琢磨了一下，浴巾也不能带进小小的淋浴间呀，不知谁设计的。没找到比其他宾馆更多的优点，560 块大洋花得有点冤，不再来了 | postive | negative |

第一个例子是一个比较明显的正向情感的表达，LSTM 之所以用错分，是由于句子中使用了过多的否定词，如“没有水”“没事”“问题”等；第二条评论陈述了商品使用前后的对比情况，给出了“千万不要去买燃气的”这样的建议，这可能误导了 LSTM，但 HB-LSTM 较好地把握了全局信息；第三条评论前后句子表达的含义出现了转折，前面夸赞了酒店的网络，但后面大段内容表达了对酒店设施、服务的不满，并表示“以后不会再来了”，因此总体来讲应该负面的情感占主导地位，对此 HB-LSTM 同样把握了总体信息，更好地考虑了总体情境在特征中的比重，作出了正确的判断。

## 3　结束语

目前在情感分析技术中一种常用的方法是将文档进行分词，去除停用词等预处理，经由一个前馈神经网络将词表示成稠密连续的词向量，再通过诸如 LSTM、CNN 等深度学习模型进行特征提取和情感分类，这种方法取得了一定的成果。本文分析了这种一般方法中存在的问题，比如情感在文档中的长期依赖问题，提出了利用文档整体所提供的情境信息来增强深度学习模型的特征提取能力，从而提高情感分类效果的假设。实验结果表明，这种整合了情境特征的模型与基线模型相比对于情感分类问题的作用是比较明显的。

在接下来的研究中将探索将情境特征的应用于更加广泛的分类问题，尤其是对社交网络文本的分析，将会在社区划分、舆情监控和分析等应方面有更大的价值。另一方面，类比于英文中字符级的深度学习模型[28]、图像处理中的基于像素点的模型，本文认为汉字是中文的基本单位，基于汉字的情感分析深度学习模型非常值得深入研究。

## 参考文献：


[1] Tsytsarau M, Palpanas T. Survey on mining subjective data on the Web [J]. Data Mining and Knowledge Discovery, 2012, 24 (3): 478-514.

[2] Pang Bo, Lee L, Vaithyanathan S. Thumbs up?Sentiment classification using machine learning techniques [C]// Proc of Conference on Empirical Methods in Natural Language Processing. 2002: 79-86.

[3] Xia Rui, Xu Feng, Zong Chengqing, *et al.* Dual Sentiment Analysis: Considering Two Sides of One Review [J]. IEEE Trans on Knowledge and Data Engineering, 2015, 27: 2120-2133.

[4] Pang Bo, Lee L. Seeing stars: exploring class relationships for sentiment categoration with respect to rating scales [C]// Proc of the 43rd Annual Meeting on Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2005, 115-124.

[5] 赵妍妍，秦兵，刘挺. 文本情感分析 [J]. 软件学报, 2010, 21 (8): 1834-1848. (Zhao Yanyan, Qin Bing, Liu Ting. Sentiment analysis [J]. Journal of Software, 2010, 21 (8): 1834−1848. )

[6] Li Shoushan, Xia Rui, Zong Chengqing, *et al.* A framework of feature selection methods for text categorizatio [C]// Proc of the 47th Annual meeting of ACL. Stroudsburg: Association for Computational Linguistics. 2009, 692-700.

[7] Wan Xiaojun. Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis [C]// Procof EMNLP. Stroudsburg: Association for Computational Linguistics. 2008, 553-561.

[8] Lin Chenghua, He Yulan, Everson R, *et al.* Weakly supervised joint sentiment-topic detection from text [J]. IEEE Trans on Knowledge and Data Engineering, 2012, 24 (6): 1134-1145.

[9] 孙艳，周学广，付伟. 基于主题情感混合模型的无监督文本情感分析 [J]. 北京大学学报：自然科学版, 2013, 49 (1): 102-108. (Sun Yan, Zhou Xueguang, Fu Wei. Unsupervised Topic and sentiment unification model for sentiment analysis [J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2013, 49 (1): 102-108. )

[10] 李素科，蒋严冰. 基于情感特征聚类的半监督情感分类 [J]. 计算机研究与发展, 2013, 50 (12): 2570-2577. (Li Suke, Jiang Yanbing. Semi-supervised sentiment classification based on sentiment feature clustering [J]. Journal of Computer Research and Development, 2013, 50 (12): 2570-2577. )

[11] 周志华. 基于分歧的半监督学习 [J]. 自动化学报, 2013, 39 (11): 1871-1878. (Zhou Zhihua. Disagreement-based Semi-supervised Learning [J]. Acta Automatica Sinica. 2013, 39 (11): 1871-1878. )

[12] 张晓梅，李茹，王斌，等. 基于融合特征的微博主客观分类方法 [J]. 中文信息学报, 2014, 28 (4): 50-57. (Zhang Xiaomei, Li Ru, Wang Bin, *et al.* Subjective and objective classification of micro-blog based on feature fusion [J]. Journal of Chinese Information Processing, 2014, 28 (4): 50-57. )

[13] Zhao Yanyan, Qin Bing, Che Wanxiang, *et al.* Appraisal expression






recognition with syntactic path for sentence sentiment classification [J]. International Journal of Computer Processing of Oriental Languages. 2011, 23 (1): 21-37.

[14] 宋艳雪，张绍武，林鸿飞. 基于语境歧义词的句子情感倾向性分析 [J]. 中文信息学报, 2012, 26 (3): 38-44. (Song Yanxue, Zhang Shaowu, Lin Hongfei. Sentence sentiment analysis based on ambiguous words [J]. Journal of Chinese Information Processing, 2012, 26 (3): 38-44. )

[15] Socher R, Perelygin A, Wu J Y, *et al.* Recursive deep models for semantic compositionality over a sentiment treebank [C]// Proc of Conference on Empirical Methods in Natural Language Processing. 2013. 1631-1642.

[16] 路斌，万小军，杨建武，等. 基于同义词词林的词汇褒贬计算 [C]// 第七届中文信息处理国际会议论文集. 2007: 17-23. (Lu Bin, Wan Xiaojun, Yang Jianwu, *et al.* Using tongyici cilin to compute word semantic polarity [C]// Proc of the 7th International Conference on Chinese Computing. 2007: 17-23. )

[17] 张磊，李梦诗，陈黎，等. 基于双层 HHMM 的产品评论特征和情感分类 [J]. 四川大学学报：工程科学版, 2013, 45 (2): 94-102. (Zhang Lei, Li Mengshi, Chen Li, *et al.* Features and opinions classification of Chinese product reviews based on two level HHMMs [J]. Journal of Sichuan University Engineering: Science Edition, 2013, 45 (2): 94-102. )

[18] Severyn A, Moschitti A. UNITN: training deep convolutional neural network for Twitter sentiment classification [C]// Proc of the 9th International Workshop on Semantic Evaluation. 2015: 464–469.

[19] Vo D, Zhang Yue. Target-dependent Twitter sentiment classification with rich automatic features [C]// Proc of International Joint Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2015: 1347–1353.

[20] Nguyen T H, Shirai K. PhraseRNN: phrase recursive neural network for aspect-based sentiment analysis [C]// Proc of Conference on Empirical Methods in Natural Language Processing. 2015: 2509–2514.

[21] Denil M, Demiraj A, Freitas N D. Extraction of salient sentences from labelled documents [EB/OL]. [2018-04-25] https://arxiv. org/pdf/1412. 6815.

[22] Li Jiwei, Luong M T, Jurafsky D. A hierarchical neural autoencoder for paragraphs and documents [C]// Proc of the 53rd Annual Meeting of the Association for Computational Linguistics. 2015: 1106–1115.

[23] Lee J Y, Dernoncourt F. Sequential short-text classification with recurrent and convolutional neural networks [C]// Proc of Conference of the North American Chapter of the Association for Computational Linguistics [C]. Stroudsburg: Association for Computational Linguistics, 2016, 515-520.

[24] Ruder S, Ghaffari P, Breslin J G. A hierarchical model of reviews for aspect-based sentiment analysis [C]// Proc of Conference on Empirical Methods in Natural Language Processing. 2016: 999–1005.

[25] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural Computation, 1997, 9 (8): 1735–1780.

[26] 数据集来源 [EB/OL] (2016-06-29) , [2018-04-25]. http://spaces. ac. cn/archives/3863/

[27] Choi H, Cho K, Bengio Y. Context-dependent word representation for neural machine translation [J]. Computer Speech & Language, 2016.

[28] Xiang Zhang, Yann LeCun, Text Understanding from Scratch [EB/OL] , (2015) [2018-04-25] https://arxiv. org/abs/1502. 01710

[29] Chen Huimin, Sun Maosong, Tu Cunchao, *et al.* Neural sentiment classification with user and product attention [C]// Proc of Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2016: 1650-1659.

[30] Tang Duyu, Qin Bing, Liu Ting. Learning semantic representations of users and products for document level sentiment classification [C]// Proc of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. 2015: 1014-1023.

[31] Ghosh S, Vinyals O, Strope B, *et al.* Contextual LSTM (CLSTM) models for Large scale NLP tasks [EB/OL]. (2016) [2018-04-25]. http://arxiv. org/abs/1602. 06291.